

**PATENT APPLICATION**

**TECHNIQUES FOR STORING MULTIMEDIA INFORMATION WITH  
SOURCE DOCUMENTS**

Inventor(s): Berna Erol, a citizen of Canada, residing at  
1422 Bellevue Avenue, Unit 105  
Burlingame, CA 94010

Jonathan Hull, a citizen of The United States, residing at  
751 Laurel Street, PMB 434  
San Carlos, CA 94070

Dar-Shyang Lee, a citizen of The United States, residing at  
32533 Carmel Way  
Union City, CA 94587

Assignee: Ricoh Company, Ltd.  
15-5 Minami Aoyama 1-Chome  
Minato-Ku  
Tokyo, Japan 107-8544

Entity: Large

## **TECHNIQUES FOR STORING MULTIMEDIA INFORMATION WITH SOURCE DOCUMENTS**

### **CROSS-REFERENCES TO RELATED APPLICATIONS**

- 5   **[0001]** The present application is a non-provisional of and claims priority from U.S. Provisional Application No. 60/462,412, Attorney Docket No. 15358-009400US, filed April 11, 2003, the entire contents of which are herein incorporated by reference for all purposes.
- [0002]** The present application incorporates by reference for all purposes the entire contents of the following:
- 10   **[0003]** U.S. Application No. 10/412,757, Attorney Docket No. Docket No. 15358-009500US, filed April 11, 2003;
- [0004]** U.S. Application No. \_\_/\_\_,\_\_, Attorney Docket No. 15358-009420US, filed concurrently with the present application;
- [0005]** U.S. Application No. \_\_/\_\_,\_\_, Attorney Docket No. 15358-009430US, filed  
15 concurrently with the present application; and
- [0006]** U.S. Application No. 10/001,895, Attorney Docket No. 15358-006500US, filed November 19, 2001.

### **BACKGROUND OF THE INVENTION**

- 20   **[0007]** The present invention generally relates to techniques for associating recorded information with a source document and more particularly to techniques for determining association information for a piece of the recorded information and information in the source document.
- [0008]** Presentations are typically given using one or more presentation slides. Slides from  
25 a source document, e.g., a Powerpoint™ (PPT) document, may be output during a presentation. The information captured or recorded during a presentation may include image information, including images of slides presented during the presentation, audio information captured during the presentation, video information captured during the presentation, and other types of information. The captured or recorded information may be stored in a

multimedia document. Multimedia documents may be considered as compound objects that comprise video, audio, closed-caption text, keyframes, presentation slides, whiteboard capture information, as well as other multimedia type objects. Examples of multimedia documents include documents storing interactive web pages, television broadcasts, videos, presentations, text documents, or the like.

[0009] If a user is interested in a previously given presentation, the user is typically limited to reviewing hard copies of the presentation slides that were used. In this case, the user cannot listen to the explanations and discussions that were given for each slide. In other cases, the presentation may be recorded in audio, video, or any combination thereof. The user can then watch the audio/video recording from start to finish while referencing the hard copy of the presentation slides. The user may view the audio/video recording of the entire presentation but really would like to refresh their memory for only certain parts of the presentation.

[0010] Often, given hardcopies of slides, a user would like to find portions of information captured during a presentation related to one or more slides of the source document. Presently, the user is limited to visually and manually searching through the audio/video recording for the presentation to identify portions of the presentation that correspond to slides that may be of interest to the user. The above method includes many disadvantages. For example, time is wasted searching for the corresponding audio/video sections. Also, a user may become frustrated at the process of searching for corresponding audio/video sections and stop searching.

## BRIEF SUMMARY OF THE INVENTION

[0011] Embodiments of the present invention relate to techniques for determining association information for recorded information and source documents.

[0012] In one embodiment, a method of storing information for recorded information and a source document is provided. The method comprises: determining a first piece of information included in the recorded information; determining at least a first source document from one or more source documents that comprises information that matches the first piece of information; and storing information identifying the recorded information, information representative of the first piece of information, and information identifying the first source document such that the information identifying the first source document can be

determined given the information identifying the recorded information and the information representative of the first piece of information.

[0013] In another embodiment, a method of storing information for recorded information and a source document is provided. The method comprises: determining information  
5 identifying a first piece of information from a source document; determining at least a first recorded information that comprises information that matches the first piece of information; determining one or more portions of the first recorded information comprising information that is associated with the a first recorded information; and storing information identifying the source document, information representative of the first piece of information, information  
10 identifying the first recorded information, and information identifying the one or more portions of the first recorded information such that the information identifying the first recorded information and the information identifying the one or more portions of the first recorded information can be determined given the information identifying the source document and the information representative of the first piece of information.

[0014] In yet another embodiment, a structure that stores information for one or more portions in a source document and recorded information is provided. The structure comprises: source document information identifying a source document; one or more source identifiers identifying the one or more portions of the source document; and information  
15 identifying one or more portions of the recorded information for each source identifier of the one or more source identifiers, wherein the one or more source identifiers and the source document information can be determined given the information identifying the one or more portions of the recorded information.  
20

[0015] In another embodiment, a structure including information that links one or more portions in a source document with one or more sections of multimedia information is  
25 provided. The structure comprises: source document information identifying a source document; a recorded information identifier identifying the recorded information; and information identifying one or more portions of the recorded information for recorded information identifier, wherein the source document information can be determined given the information identifying the recorded information identifier and the one or more portions of  
30 the recorded information.

[0016] A further understanding of the nature and the advantages of the inventions disclosed herein may be realized by reference of the remaining portions of the specification and the attached drawings.

5

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0017] Fig. 1 depicts a simplified block diagram of a system for capturing information according to one embodiment of the present invention;

[0018] Fig. 2 is a simplified block diagram of a data processing system that may incorporate an embodiment of the present invention;

10

[0019] Fig. 3 depicts a system to determine source documents and portions of source documents that comprise images included in the information stored by multimedia document according to one embodiment of the present invention;

[0020] Figs. 4A and 4B illustrate embodiments of tables that may be used to associate portions of mdoc and portions of source document;

15

[0021] Fig. 5 depicts a simplified flow chart of a method for associating recorded information with source document according to one embodiment the present invention;

[0022] Fig. 6 depicts a simplified flow chart of a method for determining slides of a source document that correspond to information in an mdoc according to one embodiment of the present invention;

20

[0023] Fig. 7 depicts a simplified flow chart of a method for using temporal information to determine a portion of a source document according to one embodiment of present invention;

[0024] Fig. 8 depicts a simplified flow chart of a method for accessing portions of an mdoc using portions of source document;

25

[0025] Fig. 9 depicts a flow chart a method for accessing a portions of a source document using portions of an mdoc according to one embodiment of the invention; and

[0026] Fig. 10 illustrates a user interface for accessing associated portions of recorded information and portions of a source document according to one embodiment of the present invention.

## DETAILED DESCRIPTION OF THE INVENTION

[0027] Fig. 1 depicts a simplified block diagram 100 of a system for capturing information according to one embodiment of the present invention. It will be apparent that system 100 as depicted in Fig. 1 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art will recognize other variations, modifications, and alternatives.

[0028] A presentation driver device 102 and a display device 104 are used to output slides and other information that may be stored in a source document 108 or a symbolic presentation file. For example, slides from a Powerpoint™ (PPT) presentation may be output and displayed on display device 104. In one embodiment, the term "source document" as used in this application is intended to refer to any document stored in electronic form. For example, a document that is created using an application program and whose contents or portions thereof may be a source document. Also, source documents may be scanned documents, a pdf version of a document, an image of a document, etc. The contents of source document 108 may include slides, images, text information, etc. A source document may comprise one or more portions. For example, a PPT document may comprise one or more pages. Each page of a PPT document may comprise one or more slides. The portions of source document 108 will be referred to as slides for discussion purposes but it will be understood that a slide may also be one or more images, one or more pages of a document, etc. Source document 108 may be created using one or more application programs. For example, a PPT document may be created using an application program, such as Microsoft's Powerpoint™. Source document 108 is an electronic document that may be manipulated and edited using the application program that created it, or any other program.

[0029] In one embodiment, source document 108 is different than a captured image of a slide, which has not been created by an application program and is often not directly editable by the application program. For example, a PPT document comprising one or more slides created using a Powerpoint™ application program can be easily edited by the Powerpoint™ application. In contrast, a joint photographics group (JPEG) image of the displayed slide is not created by the Powerpoint™ application but is recorded information. Although a PPT document may contain JPEG images, the JPEG images are included in a slide created by a PPT application.

[0030] When a slide of source document 108 is displayed on display device 104, it is referred to as an outputted slide 106. For example, outputted slide 106 is a slide or image from source document 108 that has been outputted and displayed.

5 [0031] While a presenter is giving a presentation, the presenter may display slides from source document 108 on display device 104. While a slide is being displayed on display device 104, the presenter will then often describe or explain the contents of the displayed slide. For example, the presenter may embellish on the text or images displayed in mdoc 106. Attendees of the presentation may also comment on the displayed slide (e.g., ask questions about the slide, etc.). The information output during a presentation, including  
10 information output by display device 104, by the presenter, by attendees of the presentation, or any information captured during the presentation may be captured or recorded using one or more capture devices 110. Examples of presentations include lectures, meetings, speeches, conferences, classes, demonstrations, etc.

[0032] Information recorded or captured during a presentation may include text  
15 information, graphics information, animation information, sound (audio) information, video information, slides information, whiteboard images information, and other types of information. For example, a video recording of presentation may comprise video information and/or audio information. In certain instances the video recording may also comprise close-captioned (CC) text information which comprises material related to the video information,  
20 and in many cases, is an exact representation of the speech contained in the audio portions of the video recording. Recorded information is also used to refer to information comprising one or more objects wherein the objects include information of different types. For example, objects included in recorded information may comprise text information, graphics information, animation information, sound (audio) information, video information, slides  
25 information, whiteboard images information, and other types of information.

[0033] In one embodiment, the recorded information may be stored in a multimedia document 114 in a database 115. Alternatively, the recorded information may be processed in real-time as it is captured. The term “multimedia document” as used in this application is intended to refer to any electronic storage unit (e.g., a file, a directory, etc.) that stores  
30 recorded information. Various different formats may be used to store the recorded information. These formats include various MPEG formats (e.g., MPEG 1, MPEG 2, MPEG 4, MPEG 7, etc.), MP3 format, SMIL format, HTML+TIME format, WMF (Windows Media

Format), RM (Real Media) format, Quicktime format, Shockwave format, various streaming media formats, formats being developed by the engineering community, proprietary and customary formats, and others. Examples of multimedia documents 114 include video recordings, MPEG files, news broadcast recordings, presentation recordings, recorded meetings, classroom lecture recordings, broadcast television programs, papers, or the like.

[0034] Capture device 110 is used to capture any recorded information. A single capture device may capture different types of information (e.g., a video recorder can capture both audio and video information) or a single type of information (e.g., a microphone may capture only audio information. Information may be captured using various types of capture devices such as screen capture devices 110, such as a screen capture device, a digital camera, camera, cellular phone with a camera, video recorder, audio recorder, and the like. For certain capture devices 110, such as digital cameras, cell phones with the ability to capture pictures, etc., other information related to the outputted slide may also be stored along with the captured information. Information related to the user that captured the information may also be stored.

[0035] As mentioned above, as the presentation is being given, recorded information is captured by capture device 110. For example, the recorded information may be audio and visual information of the presentation. Also, the recorded information may include captured images of the outputted slides. The information is then stored in database 115 as mdoc 114.

[0036] Temporal information may be stored for the recorded information. For example, temporal information may be stored for the captured audio and video information. If the recorded information includes images, temporal information may also be stored for the captured images. For example, the temporal information for a slide image may indicate a time when the slide was displayed by display device 104. The temporal information may be used to synchronize the various types of information included in the recorded information. For example, the temporal information may be used to synchronize audio information and video information at a time that corresponds to when outputted slides were displayed.

[0037] In one example, a slide from source document 108 was displayed from a time T1 to T2. Information from a time T1 to T2 is recorded. For example, an image of the displayed slide may be recorded and stored in mdoc 114. Also, audio information and video information from those times are recorded and stored in mdoc 114. The captured audio and video information may be associated with the captured image and the times T1 and T2. This



indicates that the audio and video information was captured at the recorded times while the captured image was being displayed. Although a range of times is stored, it should be understood that other temporal information may also be stored. For example, the time when a slide is first displayed may be stored. In this case, the range that a slide was displayed will  
5 be from the start time of slide to the start time of the next displayed slide.

**[0038]** According to an embodiment of the present invention, the stored temporal information is used to associate recorded information or portions thereof to a source document 108 or a portion of the source document, as will be described below.

**[0039]** Fig. 2 is a simplified block diagram of a data processing system 200 that may  
10 incorporate an embodiment of the present invention. As shown in Fig. 2, data processing system 200 includes at least one processor 202, which communicates with a number of peripheral devices via a bus subsystem 204. These peripheral devices may include a storage subsystem 206, comprising a memory subsystem 208 and a file storage subsystem 210, user interface input devices 212, user interface output devices 214, and a network interface  
15 subsystem 216. The input and output devices allow user interaction with data processing system 202.

**[0040]** Network interface subsystem 216 provides an interface to other computer systems, networks, and storage resources 204. The networks may include the Internet, a local area network (LAN), a wide area network (WAN), a wireless network, an intranet, a private  
20 network, a public network, a switched network, or any other suitable communication network. Network interface subsystem 216 serves as an interface for receiving data from other sources and for transmitting data to other sources from data processing system 200. For example, may receive the images to be compared via network interface subsystem 216. Embodiments of network interface subsystem 216 include an Ethernet card, a modem  
25 (telephone, satellite, cable, ISDN, etc.), (asynchronous) digital subscriber line (DSL) units, and the like.

**[0041]** User interface input devices 212 may include a keyboard, pointing devices such as a mouse, trackball, touchpad, or graphics tablet, a scanner, a barcode scanner, a touchscreen incorporated into the display, audio input devices such as voice recognition systems,  
30 microphones, and other types of input devices. In general, use of the term "input device" is intended to include all possible types of devices and ways to input information to data processing system 200.

[0042] User interface output devices 214 may include a display subsystem, a printer, a fax machine, or non-visual displays such as audio output devices. The display subsystem may be a cathode ray tube (CRT), a flat-panel device such as a liquid crystal display (LCD), or a projection device. In general, use of the term "output device" is intended to include all possible types of devices and ways to output information from data processing system 200.

[0043] Storage subsystem 206 may be configured to store the basic programming and data constructs that provide the functionality of the present invention. For example, according to an embodiment of the present invention, software modules implementing the functionality of the present invention may be stored in storage subsystem 206. These software modules may be executed by processor(s) 202. Storage subsystem 206 may also provide a repository for storing data used in accordance with the present invention. For example, the images to be compared including the input image and the set of candidate images may be stored in storage subsystem 206. Storage subsystem 206 may comprise memory subsystem 208 and file/disk storage subsystem 210.

[0044] Memory subsystem 208 may include a number of memories including a main random access memory (RAM) 218 for storage of instructions and data during program execution and a read only memory (ROM) 220 in which fixed instructions are stored. File storage subsystem 210 provides persistent (non-volatile) storage for program and data files, and may include a hard disk drive, a floppy disk drive along with associated removable media, a Compact Disk Read Only Memory (CD-ROM) drive, an optical drive, removable media cartridges, and other like storage media.

[0045] Bus subsystem 204 provides a mechanism for letting the various components and subsystems of data processing system 202 communicate with each other as intended. Although bus subsystem 204 is shown schematically as a single bus, alternative embodiments of the bus subsystem may utilize multiple busses.

[0046] Data processing system 200 can be of varying types including a personal computer, a portable computer, a workstation, a network computer, a mainframe, a kiosk, or any other data processing system. Due to the ever-changing nature of computers and networks, the description of data processing system 200 depicted in Fig. 2 is intended only as a specific example for purposes of illustrating the preferred embodiment of the computer system. Many other configurations having more or fewer components than the system depicted in Fig. 2 are possible.

[0047] Fig. 3 depicts a system 300 to determine source documents 108 and portions of source documents 108 that comprise images included in the information stored by multimedia document 114 according to one embodiment of the present invention. It will be apparent that system 300 as depicted in Fig. 3 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art will recognize other variations, modifications, and alternatives.

[0048] As shown, a multimedia document 114 is input to a source document determiner 301. In one embodiment, mdocs 114 stores recorded information. For example, multimedia document 114 may store images of slides, or a portion of the image may include a slide (i.e., the image may also comprise other information such as the picture of the presenter, etc.). Also, any other content may be input, such as text, audio, video, etc. A slide may also be present in a keyframe extracted from the recorded information. Mdocs 114 may store images of slides # 1 - N, but for discussion purposes, system 200 will be described as processing a first portion of multimedia document 114 referred to as an image #2. Although an image #2 is used, it should be understood that any content may be input. For example, text, audio, video, etc. may be input and matched to portions of source documents. It should be understood that the following techniques may be used for other portions of mdoc 114.

[0049] Source document determiner (sdd) 301 determines one or more source documents from a plurality of source documents accessible to sdd 301 that comprise at least one image included in the information stored by the input multimedia document 114. For example, source document determiner 301 determines if image #2 can be matched to a portion of one or more source documents 304, 306, 308, and 310 included in a database 302. Each source document in database 302 may include information that corresponds to image #2. For example, a source document may include content such as one or more images, slides, text information, audio information, etc. that corresponds to image #2. If image #2 includes images of slides that were displayed for a source document, the image in image #2 may be matched with a slide in any of source documents 304, 306, 308, and/or 310. For discussion purposes, source document 304 is determined to include a slide present in the information stored by multimedia document 114. It should be understood that other source documents (e.g., source documents 306, 308, and/or 310) may include slides that match portions found in mdoc 114 and the following process could also be applied to the other source documents.

**[0050]** Section determiner 307 receives information from source document determiner 301 identifying one or more source documents that that comprise at least one slide included in the information stored by the input multimedia document 114. For each source document from the one or more source documents, section determiner 307 determines one or more portions  
5 of the source document that comprise a slide included in multimedia document 114. Section determiner 307 matches image #2 with one or more portions (referred to as slides 312) of source document 304. For example, image #2 may include information that corresponds to a slide #2 in slides #1-3 in source document 304. Thus, image #2 may contain an image of slide #2 in slides 312. In one embodiment, while slide #2 was being displayed, an image of  
10 slide #2 was captured as information in multimedia document 114. As will be described below, various techniques may be used to match portions of mdoc 114 to any number of slides in slides 312.

**[0051]** Multimedia information associator (mia) 316 receives information from section determiner 307 indicating the determined slide of source document 304. Multimedia  
15 information associator 316 then determines association information for image #2 of mdoc 114 and slide #2 of source document 304. For source document 304, mia 316 determines the identity of source document 304 (e.g., filename of the source document, a portion of source document 304 (e.g., the page on which the slide is present), and a location where source document 304 is stored (e.g., directory path name). For multimedia document 114, mia 316  
20 determines information on how multimedia document 114 can be accessed (e.g., the directory or storage device identifier where multimedia document 114 is stored), when the portions of multimedia document 114 associated with an image is present (portions are identified using temporal information e.g., a starting time T1 and an end time T2). For a particular multimedia document 114, multiple portions of multimedia document 114 may comprise the  
25 slide.)

**[0052]** The above information is then stored in a table in database 320. Fig. 4A shows a table 500 where information is stored for each multimedia document (i.e., the first column identifying the multimedia document is the primary key for the table). Fig. 4B shows a table where information 550 is stored for each source document (i.e., the first column identifying  
30 the multimedia document is the primary key for the table). Mia 316 may store information for table 500 and/or table 550.

[0053] In one example, recorded information from the times T1 to T2 is associated with the image in mdoc 114. For example, the recorded information (e.g., audio/video information) may have been captured while the image was being displayed in a presentation. Recorded information from the times T1 to T2 is then associated with the determined one or more slides. Thus, recorded information that was captured while an image in mdoc 114 was being displayed is associated with a slide in source document 108 that includes information in the image of mdoc 114. Accordingly, using the image in mdoc 114, recorded information can be associated with a slide in source document 108 that includes information in the image of mdoc 114. Information for the association is then stored in database 320.

[0054] In one embodiment, table 500 includes association information that may allow a user to specify (or select) a slide in a source document and access corresponding information in mdoc 114. For example, information from a time T2 to T3 may be selected and a slide #2 in source 304 is accessed and displayed. For example, table 500 may have pre-determined images in source document 304 that correspond to temporal information for the image #2 of mdoc 114 from the times T2 to T3. Using a location of source document 304, an image in source document 304 is accessed and displayed when information from the times T2 and T3 is selected.

[0055] In another embodiment, table 550 includes association information that may allow a user to specify (or select) a portion of mdoc 114 and access corresponding information in a source document. For example, slide #2 may be selected and recorded information from mdoc 114 is accessed and displayed. For example, table 550 may have pre-determined temporal information for the image #2 of mdoc 114 from the times T2 to T3. Using a location of mdoc 114, information from the times T2 and T3 are accessed and played.

[0056] Figs. 4A and 4B illustrate embodiments of tables that may be used to associate portions of mdoc 114 and portions of source document 108. It will be apparent that tables 500 and 550 as depicted in Figs. 4A and 4B are merely illustrative of embodiments incorporating the present invention and do not limit the scope of the invention as recited in the claims. One of ordinary skill in the art will recognize other variations, modifications, and alternatives.

[0057] Referring to Fig. 4A, table 500 identifies information in mdoc 114 and information for slides in source document 108. In table 500, for each multimedia document 114 identified in a column 504, information is stored identifying one or more source documents

108 whose contents include at least one slide included in the contents of multimedia document 114.

[0058] Each row of column 504 stores information 502 identifying a multimedia document 114. A row identifies a multimedia document identifier that includes or references  
5 information. The multimedia document identifier may be a filename storing recorded information, a directory location or one or more files storing the recorded information, etc. In one embodiment, the identifier is a uniform resource locator (URL) that can be used to access recorded information, a reference to a storage location, an ID that identifies and makes the link to a non-electronic medium (e.g., paper), etc. An example describing a system that uses  
10 an ID that identifies and makes the link to a non-electronic medium is described in U.S. Application No. 10/001,895, Attorney Docket No. 15358-006500US, filed November 19, 2001, which is incorporated by reference in its entirety for all purposes.

[0059] A column 505 identifies an index type for the recorded information identified in column 504. The index type can be a number, a page number, time, range, section, line  
15 numbers, layer, etc. Although the index type is explicitly stated in column 505, the index type may be interpreted based on the recorded information identified in column 504. For example, it may be determined that a video is being referenced and the index type is a time.

[0060] Each row of a column 506 stores information identifying an index for each identifier in column 504. The index may be a range or a single pointer. For example, the index may be  
20 a temporal identifier for a portion of recorded information. Also, the index may be a pointer to a single time, one or more page numbers, etc. For example, the index may be "00:01:10 - 00:02:10", "00:09:00", "5" or "5-10". In one embodiment, each index is a start time for recorded information stored at an identifier in column 504. As shown, a first portion of recorded information starts at a time  $T = 00:00:00$  and has no duration. A second portion of  
25 recorded information is found in a byte range of "123344 - 167722". The third and fourth portions of recorded information are found at a page number "1" and a layer "4".

[0061] Each portion of recorded information is also associated with a portion of source document 108. Column 516 includes source document identifiers that identify a location of a source document 108. As shown, the location "D:\k.ppt" identifies a source document 108  
30 that is associated with recorded information in the same row. The same is true for the other rows of column 516. It will also be understood that a row in column 516 may be associated with multiple portions of recorded information in column 504.

[0062] Each row in a column 518 includes the reference to a portion of each source document 108 identified in column 516. For example, an index #3 (e.g., a slide #3) in source document 108 found at "D:\k.ppt" corresponds to temporal identifier 508. An index #1 (e.g., a slide #1) in source document 108 found at "D:\m.ppt" corresponds to temporal identifier 510 and an index #2 in the same source document 108 corresponds to temporal identifiers 512 and 514.

[0063] A column 517 identifies an index type for the source document 108 identified in column 516. The index type can be a number, a page number, time, range, section, line numbers, layer, etc. Although the index type is explicitly stated in column 517, the index type may be interpreted based on the source document 108 identified in column 516. For example, it may be determined that a PPT document is being referenced and the index type is a slide number or range.

[0064] The above table may be stored with or linked to a recorded information stream found at the URL "http://....../video.rm". Thus, when the stream is accessed at certain index points, table 500 is used to determine the corresponding portions of source document 108. The portion may then be accessed, retrieved, and/or displayed.

[0065] In Fig. 4b, table 550 identifies information in mdoc 114 and information for slides in source document 108. In table 550, for each source document 108 identified in column 551, information is stored identifying one or more mdocs 114 whose contents include at least one slide included in the contents of source document 108.

[0066] Each row of column 551 stores information identifying a source document 108. Each row identifies a source document identifier that includes or references information. The source document identifier may be a filename storing source document 108, a directory location of one or more files storing source document 108, etc. In one embodiment, the identifier is a storage location that can be used to access source document 108. Although an identifier for source document 108 is shown in each row, it will be understood that multiple rows may be associated with a single identifier.

[0067] A column 552 identifies an index type for the source document 108 identified in column 551. The index type can be a number, a page number, time, range, section, line numbers, layer, etc. Although the index type is explicitly stated in column 552, the index type may be interpreted based on the source document 108 identified in column 551. For

example, it may be determined that a PPT document is being referenced and the index type is a slide number or range.

[0068] Each row of a column 553 includes information for a portion of source document 108. For example, a portion identifier references a page, slide number, etc. in a file located at an identifier in column 551.

[0069] Each row of a column 555 stores mdoc identifiers identifying a portion of mdoc 114. Each identifier references recorded information at a location, such the URL "http://.../video.rm". In one embodiment, the URL references video content in a file "video.rm".

[0070] A column 557 identifies an index type for the recorded information identified in column 555. The index type can be a number, a page number, time, range, section, line numbers, layer, etc. Although the index type is explicitly stated in column 557, the index type may be interpreted based on the recorded information identified in column 555. For example, it may be determined that a video is being referenced and the index type is a time.

[0071] Information in a column 563 indicates which portions of an mdoc identifier correspond to information found at source identifier. For example, the video content at "http://.../video.rm" from the times  $T = 00:05:00$  to  $T = 00:09:05$  is associated with a source identifier at "D:\k.ppt". Also, video content at "http://.../video.rm" from the range "00:19:00 - 00:20:35" is associated with a portion of source document 108 at "D:\k.ppt", the whole document at "http://.../im.jpg" is associated with a portion of source document at "D:\m.ppt", and a bit at "http://.../audio.wav" at "14300" is associated with a portion of source document 108 at "D:\m.ppt"

[0072] The above table 550 may be stored with or linked to source documents 108 found at the locations "D:\k.ppt" and "D:\m.ppt". Thus, when portions of source document 108 are accessed at certain index points, table 550 is used to determine the corresponding portions of mdoc 114. The portion may then be accessed, retrieved, and/or displayed.

[0073] Accordingly, information may be selected in a source document 108 and corresponding recorded information in mdoc 114 may be accessed. For example, a slide in a presentation that was shown is selected and the video/audio information of the presenter describing the displayed slide is accessed.



[0074] In one embodiment, the input may be a source document 108, comprising one or more slides, and multimedia documents and/or portions of multimedia documents for each slide in the source document are determined. For example, for a given source document, source document determiner (sdd) 301 determines one or more multimedia documents, from  
5 a plurality of multimedia documents 114, that store at least one slide image from the source document. Section determiner 307 determines one or more portions for mdocs 114 that include the at least one slide image. Mia 316 then associates recorded information associated with the one or more portions of mdocs 114 with the at least one slide image. Information is then stored for the at least one slide and recorded information as described above.

10 [0075] In one embodiment, process of determining association information may be done manually. Also, whether the association information was determined automatically or manually, the association information may be modified.

[0076] Fig. 5 depicts a simplified flow chart 400 of a method for associating recorded information with source document 108 according to one embodiment the present invention.  
15 The method depicted in Fig. 5 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0077] In step 402, information captured during a presentation is received. The information received in step 402 may have been captured during a presentation. The  
20 information may include information of various types , e.g., audio information, video information, images, etc., and may be stored in a multimedia document 114. For example, a presentation is given by a presenter. As the presentation is given, images are displayed from source document 108. For example, slides are displayed while a presenter provides explanation and additional information about the displayed slides. The recorded information  
25 may include audio information, video information, or any other information for the presentation. Step 402 corresponds to functions performed by section document determiner 301 in Fig. 2.

[0078] In step 404, one or more source documents 108 that are associated with a multimedia document 114 are determined. For example, there may be many source  
30 documents 108 in a database. The method determines which source documents 108 include a slide that matches information in multimedia document 114. A source document 108 is

considered to be associated with information in multimedia document 114 if source document 108 comprises at least one image that is included in the multimedia document 114.

[0079] Once one or more source documents 108 are identified, in step 406, one or more portions of source document 108 that contain the information in mdoc 114 are determined.

5 Step 406 corresponds to functions performed by section determiner 307 in Fig. 2.

[0080] In one example, an image in mdoc 114 may correspond to a slide #1 in a source document 108. The slides of source document 108 that are related to an image in mdoc 114 may be determined by many methods that will be described below.

[0081] In step 408, portions of mdoc 114 are then associated with the corresponding one or  
10 more slides of source document 108. In one example, recorded information from the times T1 to T2 has been associated with the image in mdoc 114. For example, the recorded information (e.g., audio/video information) may have been captured while the image was being displayed in a presentation. Recorded information from the times T1 to T2 is then associated with the determined one or more slides in source document 108. Thus, recorded  
15 information that was captured while an image in mdoc 114 was being displayed is associated with a slide in source document 108 that includes information in the image of mdoc 114. Accordingly, using the image in mdoc 114, recorded information can be associated with a slide in source document 108 that includes information in the image of mdoc 114. Step 408 corresponds to functions performed by mia 316 in Fig. 2.

20 [0082] In one embodiment, in steps 406 and 408, the input may be a source document 108, comprising one or more slides, and multimedia documents and/or portions of multimedia documents for each slide in the source document are determined. For example, for a given source document, source document determiner (sdd) 301 determines one or more multimedia documents, from a plurality of multimedia documents 114, that store at least one slide image  
25 from the source document. Section determiner 307 determines one or more portions for mdocs 114 that include the at least one slide image. Mia 316 then associates recorded information associated with the one or more portions of mdocs 114 with the at least one slide image. Information is then stored for the at least one slide and recorded information as described above.

30 [0083] In step 410, a table is stored between the associated portions of mdoc 114 and the one or more slides of source document 108. In one embodiment, association in the table may be stored in an XML table that links the slides of source document 108 to the portions of

multimedia information 112. Alternatively, the table may associate portions of mdoc 114 with the one or more slides of source document 108. Fig. 11 illustrates an embodiment of an XML table corresponding to table 550 of Fig. 4b. The association allows a user to select either a portion of recorded information or one or more slides of source document 108 and be provided with the corresponding one or more slides of source document 108 or portions of recorded information. Thus, recorded information captured during the presentation may be accessed using the original (e.g., an electronic copy) source document 108 or the original source document 108 may be used to access the recorded information. Also, it should be understood that other information, such as portions of a paper may be linked to source document 108.

[0084] In one embodiment, the recorded information and source document 108 may be stored together. The recorded information may be embedded into source document 108 or vice versa. The table may be stored with both the recorded information and source document 108, or separately. In the case of a Powerpoint™ file, for example, the recorded information (or the association information) is stored in source document 108 as a PPT object. The association between the slide number and the recorded information is stored in a PPT format.

[0085] If recorded information is included in source document 108 (or vice versa), a self contained document formed that allows either the recorded information or information in source document 108 to be accessed using the self-contained document. This makes it easier to share, publish, e-mail, and edit information in the recorded information or source document 108. Even if the recorded information and source document 108 are included in a self contained document, they can be separated. For example, when a user is editing the PPT file with a video file attached to it, the user can extract the video file and share it with others.

[0086] Although the processing depicted in Fig. 5 was described for one portion of an mdoc 114, it will be understood that the processing may be repeated for multiple source documents 108. Also, if source documents are inputted instead of mdocs 114, the process may be repeated for multiple of mdocs 114.

[0087] Figs. 6 and 7 depict embodiments of determining step 406 and associating step 408 of Fig. 4. The methods depicted in Figs. 6 and 7 are merely illustrative of embodiments incorporating the present invention and do not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0088] Fig. 6 depicts a simplified flow chart 600 of a method for determining slides of a source document 108 that correspond to information in an mdoc 114 according to one embodiment of the present invention. The method is described using image #2 and source document 304 of Fig. 3. Although the method is describes as matching images, it should be understood that any content may be matched. For example, content such as text, audio, video, etc. may be matched. In step 602, an image #2 is determined. Although the process will be described for only one mdoc 114, it should be understood that multiple mdocs 114 may be processed as follows.

[0089] In step 604, image #2 is matched with information in source document 304. For example, the information may be an image. In one embodiment, image matching techniques disclosed in U.S. Patent Application No. 10/412757, filed on April 11, 2003 entitled "Automated Techniques for Comparing Contents of Images" may be used and is hereby incorporated by reference for all purposes. An input image is matched to a set of potential candidate images. The set of candidate images may be the slides in the source document 304. For each multimedia document 114, individual images that are included in the information stored by the multimedia document 114 and contain slides are used as an input image. The individual images may be determined from analyzing keyframes extracted from video information, from images included in the recorded information stored by multimedia document 114, slide information, etc. The time range for each slide image may also be determined (i.e., the information shown in the fourth column of the table in Fig. 4B).

[0090] In step 606, an identifier for the matched slide in source document 302 is determined. For example, the identifier may be a slide number or page number. The identifier is used as an index into source document 302.

[0091] In step 608, a location where source document 302 can be accessed is determined. For example, the location may be a file name, URL, storage location, etc.

[0092] In step 610, information related to the identifier and location of source document 302 and the portion of multimedia document 114 including the slide are stored such that an association is formed where recorded information in mdoc 114 is associated with the identifier and location of source document 302. For example, recorded information from the times T2 to T3 is associated with the identifier of source document 302.

[0093] Fig. 7 depicts a simplified flow chart 700 of a method for using temporal information to determine a portion of a source document according to one embodiment of

present invention. The method will be described using image #2 and source document 304 of Fig. 3.

[0094] In step 702, image #2 is received. In step 704, temporal information associated with image #2 is determined. The temporal information may be time stamps associated with a portion of multimedia document 114. For example, a time stamp may have been associated with a captured image when a picture was taken or a slide was captured.

[0095] In step 706, a slide in source document 302 that corresponds to the temporal information is determined. For example, the times that slides in source document 302 were displayed may correspond to times associated with recorded information. If the temporal information indicates a time stamp, then the corresponding slide in source document 302 that was displayed at that time is determined. This can be determined using time stamps determined by a presenter. The time stamps may be expected durations for times that slides will be displayed. If a presentation is started at a certain time, the expected durations can be used to association recorded information with slides in source document 108. Also, times that slides in source document 108 are displayed may be recorded and stored. These times are used to associated recorded information with slides in source document 108.

[0096] Once the temporal information and slide in source document 302 are known, a range of times for the recorded information may be determined. If the temporal information indicates that the slide was displayed from the times between T2 and T3 and a time stamp is in between the times T2 and T3, that slide is determined.

[0097] In step 708, recorded information corresponding to the temporal information is then associated with the determined slide of source document 304. For example, recorded information from the times T2 to T3 is associated with slide # 2.

[0098] Once information in mdocs 114 and source documents have been associated, they can be used to access each other. The following describes techniques for accessing portions of mdocs 114 and source documents 108. Fig. 8 depicts a simplified flow chart 800 of a method for accessing portions of mdoc 114 using portions of source document 108. The method depicted in Fig. 8 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

**[0099]** In step 802, an input is received identifying a slide from a source document 108. The information identifies a source document 108 and index information (e.g., a slide identifier). The input may be received as a selection of a slide from an icon on a screen, an input of a slide identifier such as a page or slide number or the index number that has been stored in table 318.

**[0100]** In step 804, in response to receiving the input, a portion of mdoc 114 is determined. Based upon the received information, a table (such as table 550 depicted in Fig. 5B) can be searched to determine if the slide was shown during a captured presentation. In searching table 550 it is determined if column 551 of table 550 contains an entry that matches the information identifying the source document. If such an entry is found, search the rows in column 553 associated with the entry to determine an entry matches the particular slide received in step 802. If such an entry exists, then multimedia documents 114 corresponding to that entry and identified in column 555 are the multimedia documents that contain that particular slide. Column 563 of table 550 identifies the portions (time ranges) of multimedia documents 114. The identified portions may be recorded information.

**[0101]** In step 806, the determined portions of mdoc 114 in step 804 are accessed. For example, table 550 may cause the recorded information associated with the identifiers 556, 558, and 560 to be retrieved at the locations indicated by the identifiers. For example, video, images, and text are retrieved for recorded information at the locations shown.

**[0102]** In step 808, the determined portions of mdoc 114 associated with the slide are displayed, played, etc. In one embodiment, the portions may not be played right away but an icon or any other device that may trigger the playing of portions are displayed on a screen. When the icon is selected, the portions will be played. For example, when input is received for the slide, corresponding video information associated with identifier 556 is played in one window. The video may be a video of a person giving the presentation and may also include audio of any explanations or discussions about the slide. Also, a JPEG image associated with identifier 558 of a captured slide might be displayed in a second window and text associated with identifier 560 may be displayed in a third window. The text may be a transcription of audio information for the explanation recorded during the presentation, video paper, or the like. Various different techniques using one or more output devices may be used to output the information.

[0103] Accordingly, input was received for a slide in a source document, and corresponding portions of an mdoc 114 (e.g., recorded information) associated with the slide was automatically accessed for a user. The corresponding recorded information may then be played or displayed. Thus, users interested in a slide in a presentation may easily view recorded information that is associated with the slide.

[0104] Fig. 9 depicts a flow chart 900 a method for accessing a portions of a source document 108 using portions of an mdoc 114 according to one embodiment of the invention. The method depicted in Fig. 9 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art would recognize other variations, modifications, and alternatives.

[0105] In step 902, an input is received identifying an mdoc 114 and a portion (or index) of mdoc 114. For example, a time T1 for a video of a presentation that was given may be input. In another embodiment, a timeline may be displayed and a certain time T1 on the timeline may be selected. Additionally, an index of different time ranges may also be displayed for selection by a user.

[0106] In step 904, in response to receiving the input, a portion of source document 108 is determined. Based upon the received information, a table (such as table 500 depicted in Fig. 4A) can be searched to determine the portion of mdoc 114 was recorded while the slide was being displayed in a presentation. In searching table 500 it is determined if column 504 of table 500 contains an entry that matches the information identifying mdoc 114. If such an entry is found, the rows in column 506 associated with the entry are searched to determine an entry that matches the particular portion of mdoc 114 received in step 902. If such an entry exists, then source documents 108 corresponding to that entry and identified in columns 516 and 520 are the source documents that contain that particular portion. Columns 518 and 522 of table 500 identify the portions (slide numbers, indexes) of source documents 108. The identified portions may be images of the slides.

[0107] In step 906, the slides corresponding to the indexes of source document 108 are accessed. In one example, if the portion of mdoc 114 from the time T =5:00 to T=8:00 is input, slide #1 from the location D:\m.ppt is accessed.

[0108] In step 908, the slides corresponding to the portion of multimedia information inputted are displayed. In this case, a user may view slides that correspond to the portion of mdoc 114. Although it is described that the slide is displayed, it should be recognized that it

may not be displayed but any other action may be performed, such as storing the slide in a different location, opening the slide in an application such as Powerpoint™, sending the slide to a recipient, etc. If the slide is opened in Powerpoint™, then the slide may be edited by a user. The advantage of this is that a user having access to mdoc 114 of the presentation may  
5 access the original (or an electronic copy of) source document 108 rather than using a captured image, which would not allow editing in the same way as the original slides of source document 108 allow. This is because captured multimedia documents 114 are often images such as JPEGs and MPEGs that cannot be edited by a Powerpoint™ application while the original slides are in Powerpoint™ format that can be edited using a Powerpoint™  
10 application.

**[0109]** Accordingly, users can use a portion of multimedia information 112 to access the original slides of a source document 108 that were used to give a presentation. Thus, slides may be accessed automatically and any actions may be performed, such as displaying the slide or editing the slide.

**[0110]** In one embodiment, for each slide in a particular source document 108, association information to portions from one or more multimedia documents 114 comprising the slide may be stored in source document 108 for the slide. Various ways of association may be used, e.g., the association information may be stored in an XML file that is associated with the source file. In one embodiment, this does not affect the ability of the file to be  
20 opened/edited using the application program that was used to create the source document. Advantages of such a feature are, for example, the file can be shared, copied, etc. Also, the portions of mdocs 114 can be accessed using the application that created slides in a source document 108. For example, a Powerpoint™ application may be used to play back recorded information associated with the slides of a Powerpoint™ document. A user may also access  
25 slides in a source document while using an application configured to access portions of mdoc 114. For example, a media player may be used to access a slide in source document 108 while listening to recorded information describing the slide.

**[0111]** Fig. 10 illustrates a user interface 1000 for accessing associating portions of recorded information and from a source document 108 according to one embodiment of the  
30 present invention. It will be apparent that interface 1000 as depicted in Fig. 10 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art will recognize other



variations, modifications, and alternatives. As shown in Fig. 10, a PPT source document 108 has been opened using an application program (i.e., Powerpoint™) that was used to create source document 108.

[0112] As shown, icons 1002 associated with portions of a source document 108 are displayed along the left-hand side of interface 1000. Also, icons associated with portions of mdocs 114 may also be shown. Icons included can be selected by a user in order to display the slide in source document 108 or access portions of mdoc 114 that correspond to the slide.

[0113] A slide 1004 from source document 108 is currently being displayed in window 1005 of interface 1000. Also, portions of mdoc 114 are being displayed in thumbnails 1006 and 1008 (recorded information segment 1 and recorded information segment 2). In one embodiment, thumbnails 1006 and 1008 are static representations, i.e., the recorded information is played only upon selection of 1006 or 1008. The portions have been associated with slide 1004. In one embodiment, as slide 1004 is being displayed, either recorded information segment 1 or recorded information segment 2 is played. A segment of recorded information may be a portion of information (e.g., audio/video information) from a first time to a second time. Also, an input may be received to determine when and whether recorded information segment 1 or 2 may be played. For example, segment 1 may be played followed by segment 2 or a user may select which segment is played.

[0114] In one embodiment, recorded information segment 1 and recorded information segment 2 are part of the information that has been captured during a presentation in which the slide from source document 108 depicted in window 1004 was presented or outputted. The information may include an explanation or discussion that was recorded while slide 1004 in source document 108 was displayed. Using interface 1000, a user can view the explanation (e.g., segment 1 and 2) that was provided during the presentation for slide 1004 by accessing slide 1004, etc.

[0115] In another embodiment, recorded information segment 1 or recorded information segment 2 may be used to access a slide in source document 1004. For example, segment 1 or segment 2 is selected and a corresponding slide may be displayed. For example, segment 1 may correspond to a first slide and segment two may correspond to a second slide in source document 108. When either of the segments is selected, a corresponding slide is displayed and the recorded information is played.

[0116] While the present invention has been described using a particular combination of hardware and software implemented in the form of control logic, it should be recognized that other combinations of hardware and software are also within the scope of the present invention. The present invention may be implemented only in hardware, or only in software,  
5 or using combinations thereof.

[0117] The above description is illustrative but not restrictive. Many variations of the invention will become apparent to those skilled in the art upon review of the disclosure. The scope of the invention should, therefore, be determined not with reference to the above description, but instead should be determined with reference to the pending claims along with  
10 their full scope or equivalents.